Do economic evaluations in primary care prevention and the management of hypertension conform to good practice guidelines? A systematic review

Barton, Pelham; Jowett, Sue; Sutton, Andrew; Penaloza Ramos, Maria

DOI:
10.1177/2381468316671724

Do economic evaluations in primary care prevention and the management of hypertension conform to good practice guidelines? A systematic review

Maria Cristina Peñaloza Ramos, MA [1*], Pelham Barton, PhD [1,], Sue Jowett, PhD [1], Andrew John Sutton, PhD [2]

[1] Health Economics Unit, University of Birmingham, UK

[2] Health Economics Unit, Leeds Institute of Health Sciences, University of Leeds, UK

Key words: decision-analytic modelling, modelling, guidelines, good practice, cardiovascular disease

*Corresponding author

Health Economics Unit, Public Health Building, University of Birmingham, Edgbaston, Birmingham B15 2TT. United Kingdom, email: m.c.penaloza@bham.ac.uk Tel. +44 (0)121 414 7061; Fax: +44 (0)121 414 8969

Running title: A systematic review of applied DAM studies

Acknowledgements: We want to thank Professor Richard McManus and Professor Jonathan Mant for their insightful comments during the design of this study.

Contributors: MCP undertook the review and analyses and wrote the first draft. All authors contributed to refined manuscript and approved the final version.

Competing interests: None declared. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Health Economics Unit or the University of Birmingham. No funding was received for this manuscript.

NOTE: Underlying research material(s) related to this paper can be accessed by direct request to the corresponding author.

Abstract

Background: Results of previous research have identified the need for further investigation into the compliance with good practice guidelines for current decision-analytic modelling (DAM).

Objective: To identify the extent to which recent model-based economic evaluations of interventions focused on lowering the blood pressure (BP) of patients with hypertension conform to published guidelines for DAM in healthcare using a five-dimension framework developed to assess compliance to DAM guidelines.

Methods: A systematic review of English language articles was undertaken to identify published model-based economic evaluations that examined interventions aimed at lowering BP. The review covered the period January 2000 to March 2015 and included the following electronic bibliographic databases: EMBASE and Medline via Ovid interface and the Centre for Reviews and Dissemination's (CRD) NHS-EED. Data were extracted based on different components of good practice across five dimensions utilizing a framework to assess compliance to DAM guidelines.

Results: Thirteen papers were included in this review. The review found limited compliance to good practice DAM guidelines which was most frequently justified by the lack of data.

Conclusions: The assessment of structural uncertainty cannot yet be considered common practice in primary prevention and management of hypertension and researchers seem to face difficulties with identifying sources of structural uncertainty and then handling them

correctly. Additional guidelines are needed to aid researchers in identifying and managing sources of potential structural uncertainty. Adherence to guidelines is not always possible and it does pose challenges, in particular when there are limitations due to data availability that restrict, for example, a validation process.

Words: 254

Introduction

Cardiovascular disease (CVD), which incorporates coronary heart disease (CHD) and stroke, is the main cause of death worldwide[1] and in England and Wales.[2] Hypertension, defined as a persistent raised blood pressure (BP) of 140/90 mmHg,[3] has been recognised as the most important modifiable risk factor for CVD.[2, 3] Poorly controlled high BP can damage artery walls and increase the risk of developing a blood clot. Moreover, if it is not treated it can also damage organs such as the kidneys, heart and brain. Decision-analytic modelling (DAM) guidelines have recognised that randomised controlled clinical trials (RCTs) are good sources of evidence to judge the effectiveness of treatments; however, because the time horizon for trials often does not reflect the full duration of the impact of interventions, DAM is used to extend the results of a short term trial over a longer time horizon.[4, 5] A primary outcome used in RCTs that are focused on hypertension is often change in BP. However, this is only an intermediate outcome and DAM can be used to examine the impact of change in BP on the risk of CVD events in the longer term.

Previous research has identified the need for further investigation into the compliance of DAM to good practice and its impact on the conclusions drawn from economic evaluations.[6] Our aim is to critically evaluate how DAM in primary prevention of CVD conforms to guidelines and, in doing so, validate a framework previously developed to assess compliance to guidelines. The focus here is on one particular clinical area since this makes it possible to remove some of the variation between models which is not relevant for the purpose of assessing compliance (for example, different outcomes, treatment options or sources of uncertainty). CVD prevention has been selected due to the wide number of recent and available model-based cost-effectiveness (CE) studies conducted in this topic area. We focused on interventions aimed at lowering BP, as a modifiable risk factor for CVD, and

sought to answer the research question: 'to what extent do model-based economic evaluations of primary prevention interventions aimed at lowering BP in patients with hypertension or at risk of developing hypertension conform to the published guidelines for DAM?'

Methods

Studies of interventions aimed at lowering BP were reviewed and the challenges faced when applying DAM methods were identified and discussed. A systematic review was conducted, meeting the UK Centre for Review and Dissemination guidance and the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines for reporting.[7]

The review followed a structured approach for framing research questions: patient population (P), intervention (I), the comparator group (C), outcome (O) and the study design (S), or PICOS.[7] Papers published from January 2000 to March 2015 and written in English were included in this review if they met all of the following conditions:

- the target population was individuals presenting with high BP or at risk of developing hypertension;
- the intervention(s) aimed at lowering BP;
- management of hypertension, as a modifiable risk factor for CVD, was part of a primary prevention strategy (when studies also included secondary prevention, we have concentrated on the results for primary prevention); and
- the study was a model-based economic evaluation.

This review excluded systematic reviews, guidelines, trials, protocols and conference abstracts. In addition, we also excluded studies where the interventions:

- were aimed at screening BP;

- were part of a polypill strategy;
- measured non-adherence to treatment; or
- were part of a secondary prevention and treatment strategy.

Searches were undertaken using terms identified by expert clinical opinion and a list of synonyms identified for each term that helped inform the final search terms used in this review ("cost effectiveness", "mathematical model", "decision analysis", "Markov model", "decision tree", "economic evaluation", "hypertension" and "lowering blood pressure"). The search was undertaken using truncations and wildcards and all synonyms were subsequently combined with appropriate medical subject heading terms (MeSH) or subject terms using Boolean operators (Appendices 1 and 2).

The following databases were searched: EMBASE and Medline via the Ovid interface, and the Centre for Reviews and Dissemination's (CRD) NHS Economic Evaluation Database (NHS-EED). In addition, we manually examined the reference lists of the studies included in this review. All papers identified by database searching were exported into ENDNOTE-X7^{TM} and duplicate references were removed.

Titles identified by the searches were screened by reading the abstract; this activity was completed by two reviewers (SJ and CP). Articles that appeared to be relevant at this point were obtained and screened against the inclusion and exclusion criteria (CP); several papers appeared relevant on reading the abstract but were subsequently excluded after reading the full paper.

All studies were manually searched and data were extracted by a single reviewer (CP); any doubtful point(s) were checked with at least one another reviewer.[6] The extraction tool consisted of a framework[6] that synthesizes contemporary DAM guidelines in a single

checklist instrument; this framework was developed to aid researchers assessing adherence to guidelines. The tool aided the retrieval and organisation of information from each study across five dimensions (Appendix 3):

i) problem concept;

ii) model concept;

iii) synthesis of evidence;

iv) analysis of uncertainty; and

v) model transparency and validation.

This approach ensured that the review did not miss any information related to the model building process. Data were extracted as free text and in the form of a 'yes/no' response.

Results

The database search yielded 2607 studies; after removing 27 duplicates, 2580 studies were left for screening. 2549 studies were excluded because they did not consider a CVD related intervention, were not a model-based economic evaluation, or were focused on screening (Figure 1). 31 full-text articles were assessed for eligibility, of which 18 were rejected as a secondary prevention strategy. 13 studies were included in this review, none of which were identified through other sources (Figure 1).

Only two of the studies included were published prior to 2004. Thus it can be seen that the majority of studies (11/13) would have had access to DAM guidelines at the time of their publication (for example, Weinstein (2003)[8] or Philips (2004)[9]).

Four studies evaluated programmes for the clinical prevention and treatment of hypertension[10-13] and nine evaluated antihypertensive drug treatments to lower BP (Table

1).[14-22] Ten studies were cost-utility analyses (CUA) or combined both CUA and a cost-effectiveness analysis (CEA)[10-12, 15-17, 19-22] while three studies were CEA[13, 14, 18] (Table 1). The intervention target (risk factor) examined was high BP. The remainder of this section describes the main findings.

Problem concept and model concept

The decision problem and study objective(s) were stated in all the studies (Table 2), and all evaluated CE from a health care payer perspective. The target decision-maker audience was made explicit in 10/13 studies as that of the health care payer, i.e. including only the health effects experienced by patients receiving the intervention and costs for the medical services required to provide the intervention.[23] For the remaining studies[10, 14, 19] the perspective was left implicit. Ekman[19] commented that the analysis was "in a Swedish health-care setting", while Stevanovic[14] was interested "in the Dutch setting" and Gandjour[10] focused on those "insured by the German SHI", where SHI refers to the German Statutory Health Insurance.

For all studies, the target population was individuals with hypertension or at risk of developing hypertension (Tables 1 and 2), frequently stratified by gender, presence of hypertension, age groups, and mean age. The target population was always modelled as closed (reflecting members entering only at the start of the analysis).

Despite all the studies sharing a common aim, namely primary prevention of CVD via lowering BP, these economic models compared a wide range of interventions and presented their results using outcome measures such as QALYs,[10-12, 15-17, 19-22] life years gained (LYG),[13, 14, 16, 18, 20, 22] net health benefits (NHB),[18] net monetary benefits (NMB)[17] and expected value of perfect information (EVPI)[17] (Table 1).

Side effects were modelled in only one study.[15] Four studies[10, 11, 14, 22] acknowledged the lack of adverse events as a limitation of their results due to lack of data. Two studies argued that since 'previous clinical trials found that first-line hypertensive drugs do not have more side effects than placebo'[13] or they have 'mild side effects'[19] there was no need to model adverse effects. Similarly another study argued that fatal side effects would have been already captured in the clinical trials via the measure of effectiveness.[18]

All the studies commented on the reasons for the selection of their comparators, where their choice of comparators seems to have been governed by the scope of the study. Two studies acknowledged as a limitation the exclusion of relevant comparator(s) arguing that there may be more relevant comparators not included.[19, 21] Furthermore, the 'do nothing' option was considered in four of the studies.[10, 14, 18, 19]

All the studies used Markov models and included a figure showing the model structure; in one study[13] the structure of the Markov model shown in the figure did not seem to reflect the structure of the model described in the text. The model structures accounted for both acute and chronic health states. Five studies made explicit reference to how the structure of their models was defined either by using an existing generic model,[18] being based on disease progression[10, 11] or consisting of health states designed to reflect the course and history of CVD events.[22] One study reported that 'health states in the Markov model are based on cardiovascular events measured in the previously reported registry study'.[21] For the remaining studies it was inferred that the model structure was based on disease progression.

A lifetime time horizon was adopted in all but two studies: of these, one used a five-year[15] time horizon for a population aged 65 years whilst the second used 20-years for a population aged 18 and over.[22] The five-year time horizon was justified as matching the five-year time span given to social security authorities in China for budget planning[15] whilst the 20-year

time horizon was not discussed.[22] Cycle length, though rarely justified in the studies, was always 1 year. Only one study[10] justified their choice as most of the data used in their model referred to a 1-year period.

Synthesis of evidence

Patient heterogeneity was considered in most of the studies; results were presented by age cohorts[10, 12, 14, 15, 18] and gender.[10-12, 14-16, 18, 19, 21] Some studies added further analyses based on the risk of CVD,[10, 12, 17] scenarios of SBP reduction[14, 19] smoking[14] and patient adherence.[14, 22] The risks of secondary events were modelled in seven of the studies - e.g. the risk of a further stroke after a first stroke-.[12-14, 18-21] In some instances, assumptions were acknowledged; for example, the study by Stevanovic[14] assumed the risk of secondary events to be equal to the risk of a first non-fatal CVD event. The authors acknowledged that this would lead to an under-estimation of the CVD risk, and so an increased risk of death in patients experiencing non-fatal CVD events was adopted.[14] In Wisloff,[18] secondary non-fatal events were allowed, and a patient experiencing a secondary event was assumed to be in a health state which was worse than the state they were already in. For example, a patient with stroke sequelae that experiences a MI will have the risk and costs associated with the stroke sequelae and not those related to MI). Perman[13] utilized expert opinion in the assessment of the risk of secondary events. Montgomery,[12] due to a lack of data assumed that any second cardiovascular event was fatal and acknowledged this as a limitation. Some studies that did not use separate states to model secondary events[10, 11, 22] captured the increased mortality from secondary events through the mortality rate of patients surviving CVD events. Few of the studies acknowledged the lack of epidemiological data to model secondary events as a limitation.[11, 22]

All studies applied discounting to their results: a discount rate of 3% was most common for costs and benefits;[10, 15, 17, 19, 21, 22] two studies used a different discount rate for costs and benefits (Stevanovic used 4% and 1.5%[14] while Montgomery used 6% and 1.5%[12] respectively) (Table 1). Information on the parameters used as inputs were most frequently presented in tables showing mean values and the type of distribution(s) while some studies also included 95% confidence intervals or range intervals.[10, 11, 20] The methods used to report the sources of information varied from reporting a detailed list of sources per parameter in a table to mentioning the sources of data in the main text.

Analysis of uncertainty

The studies examined and reported uncertainty surrounding their identified outcomes through sensitivity analysis (SA). Uncertainty in parameter estimates was most commonly handled through deterministic sensitivity analysis (DSA) and probabilistic sensitivity analysis (PSA). Five studies used only one-way DSA,[12, 15, 19, 20, 22] whilst another four[11, 16-18] only used PSA. Only one study measured EVPI[17] (Table 2).

Elements pertaining to structural uncertainty (SU) were acknowledged as such in six studies.[10, 11, 13, 14, 19, 20] Most commonly SU was assessed through SA by varying the time horizon,[11] the duration of the effectiveness of the treatment,[11, 14] the discount rate[13, 19] or by using alternative measures of outcomes.[19] One study examined the impact of assumptions related to secondary events.[18] Lack of clinical evidence for key parameters such as the treatment effect of drugs[10, 11, 14] was identified as a source of SU. Two studies acknowledged that they could have included more relevant comparators had they had more information,[19, 21] and another two acknowledged that they had excluded a potentially relevant state due to lack of epidemiological data[10] or insufficient evidence on its relevance.[20]

The decision about which events and health states were included was partially discussed. Some studies acknowledged that they subdivided a health state[11] (CHD into MI, HF and angina), or excluded a potentially relevant health state[15] (combined stroke and MI event). All studies included chronic health states (post events); however, few discussed having modelled the progression of disease.[10, 11, 22] Most frequently, the studies acknowledged the adoption of assumptions, i.e. assuming the duration of treatment effects to be lifetime or as long as the time horizon in the model,[10, 11, 14] or five years[19] or varied.[21]

Model transparency and validation

All the studies included a graphical description of the Markov model they used (Table 2). Sources of funding were identified in 11 studies: five were funded by the pharmaceutical industry,[14-16, 19, 21, 22] one benefited from joint funds from government and pharmaceutical sources, [18] three were exclusively government-funded,[11, 12, 20] and one was privately funded.[17] None of the studies stated any means for accessing more detailed information about the model. All the studies had a clear policy context with an explicit statement of funder and developer.

Validation, according to guidelines,[24] is a set of methods for judging the accuracy of a model in making relevant predictions, in other words, validation helps readers understand what a model does and how it does it. In this review we checked for five main types of validation. All the studies were subjected to face validity checks (having been peer reviewed and published in a journal) and they were subjected to verification (internal validity checking). The methods used were justified to a greater or lesser extent in each study. All studies undertook SA of parameters as a way to double check that the direction and magnitude of their outputs were as expected.

In terms of cross validation, results were mixed. Eight studies[10-14, 17, 19, 20] examined different models that addressed the same problem and compared their results; however the level of detail provided varied. Five studies presented limited or no evidence of cross-validation;[15, 16, 18, 21, 22] only Wisloff[18] undertook an exercise of external validation by comparing their estimated lifetimes to those reported by Statistics Norway and in doing so they found that the input into their model needed to be adjusted to fit Norwegian mortality data. An assessment of predictive validity was not included in any of the studies considered.

Discussion

Using a previously developed practical framework,[6] we have critically evaluated how 13 published economic evaluations conformed to contemporaneous good practice guidelines. We found that published economic evaluations of interventions aimed at lowering BP in patients with hypertension, as part of a primary prevention strategy of CVD, demonstrate limited compliance to DAM guidelines which has usually been explained by lack of data or imperfect data. This was particularly apparent in the assessment of SU (or lack of) and model external validation.

This review identified common grounds in terms of the adherence to, and use of, guidelines. The conceptual modelling in all the studies included in this review was based on a disease process where the focus was on the definition of the health states (conditions) as opposed to treatment (pathways) received and where the decision problem posed required the evaluation of the reduction in the risk of developing hypertension, thus explaining the use of Markov models.

It has been argued that alternative model structures can lead to variations in model predictions,[25] most importantly, in the context of a primary prevention strategy, an inappropriate model structure may lead to poorly informed policy decisions, resulting in inefficient allocation of scarce resources.[26] Models are by nature sensitive to choices made at every single stage during the model development process (i.e., model concept, model structure). There will almost always be more than one set of choices, for this reason, guidelines have suggested assessing the extent to which model predictions are influenced by the choices made during the model development process, and have suggested methods to do so, such as scenario analyses.[27, 28]

Lifetime time horizons should be adopted (or be justified when constrained by the cohort's lifetime) or at the very least, time horizons should be 'long enough' to capture relevant differences in outcomes across strategies.[23] Lack of data or imperfect data still poses important challenges for researchers - for example, when modelling the risk of secondary events and disease progression or to attempt the assessment of model validity -. Even though elements pertaining to SU were identified by various authors, the assessment of SU cannot be considered common practice in this particular clinical area and additional guidelines are still needed to aid researchers identifying and quantifying SU.

External validity still poses a challenge to researchers and more importantly, to future guidelines due to the apparent unavailability of actual extra data (from RCT or patient level data) to undertake the exercise. It has been suggested that instead of using all the data available to create a model, some data be set aside to use during the validation process (for example, one-third of the data).[29] This may or may not always be possible, and will depend on how much data a researcher has to build a model.

Studies included in this review shared similar research questions and yet there was a great diversity in the structures of the Markov models used. Some of these were simple and some more complex, and they were generally developed with limited justification.[26] These indicate, as suggested by Squires, et al (2016),[30] that the methods for the development of the model structure are still underdeveloped. This can lead to errors including poor validity, credibility, and no basis for model verification and the analysis of structural uncertainty.

Caro and Möller[29] described the above as the disposable approach to modelling: models are built for a single use, focused on a particular product for a relatively short time. This explains - to some extent - the reduced motivation for undertaking model validation.[29] Future research should examine whether the development of 'generic models', or, as proposed by Caro and Möller, the development of multi-use models over time, can capture sufficient detail to be realistic and avoid particulars for which there are no data, and thereby allow the economic evaluation of interventions targeting CVD in any setting, and whether this will bridge the knowledge gap and, most importantly, allow ease of comparison between the results obtained from different studies.

This is the first study that has critically reviewed compliance to DAM guidelines using a previously developed practical framework. It has covered more than a decade of published DAM studies of interventions aimed at lowering BP in patients with hypertension. We believe the inclusion of recent studies from European, American and Asian countries has helped to reflect current practice worldwide.

The exclusion criteria adopted may be considered as limitation; however, these were required to guarantee consistency in the analysis. Furthermore, a negligible number of non-English-language studies were identified pertaining to applied studies. The fact that none of the studies included was published after the release of the 'five-dimension framework' and the

selection of one particular clinical area (and any impact on generalisability this may have) may also be considered a limitation.

Our findings seem in line with recent debate around the methodological challenges being faced by DAM where model validation and SU have been identified as fundamental problems due to the lack of motivation, time and data to validate models and, in the case of SU, a lack of methods.[29]

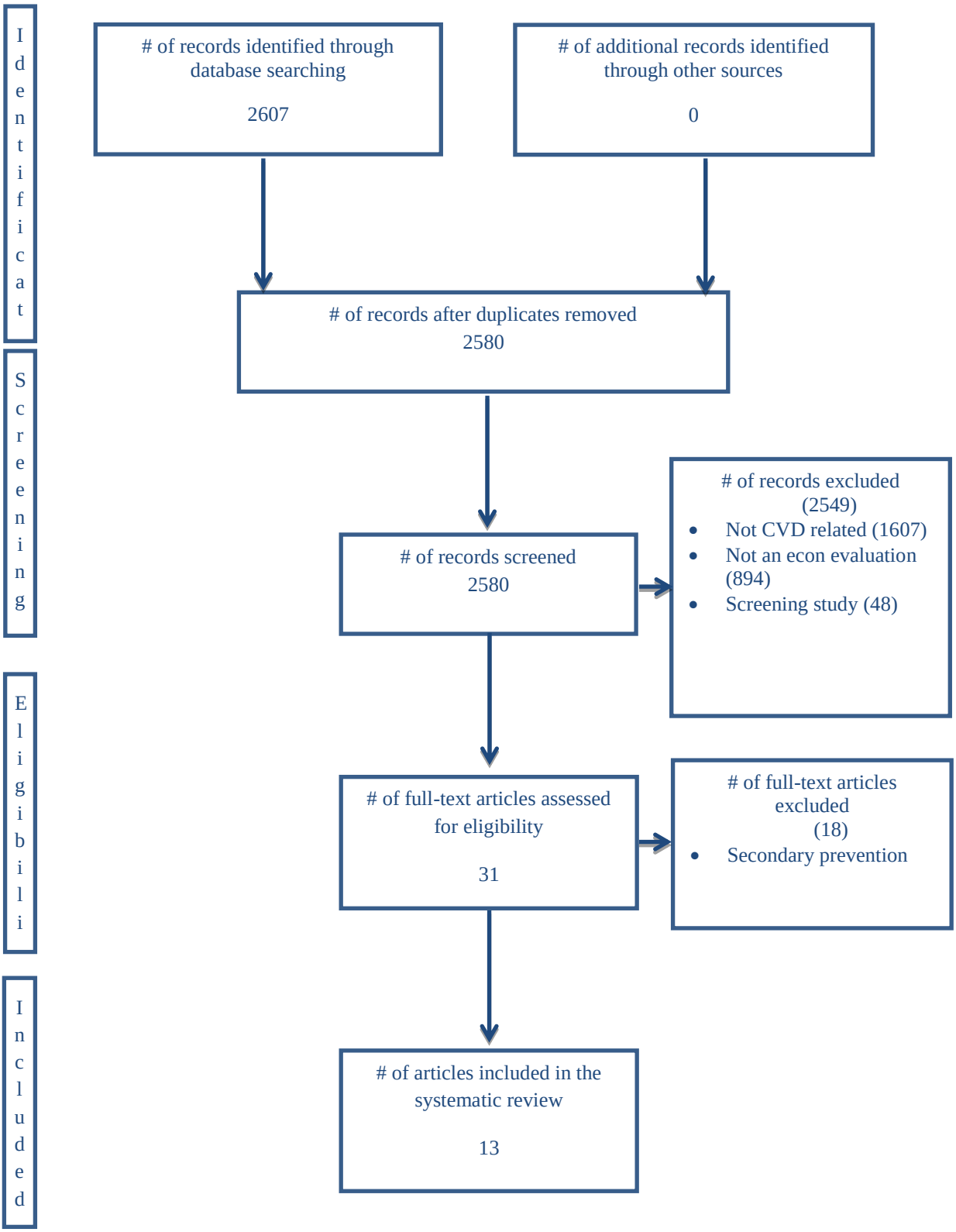

Figure 1 Flow chart

Table 1 Summary of analytic framework, methods and model features of studies included

Study/ Year	Research Question	Perspective/ Country	Comparators	Target Population/ Intervention	Time Horizon/ Discount Rate	Reported Baseline Results	Health States	Study Design and Outcomes
Kaambwa et al, 2014 [11]	LT CE of self-management of HPN	UK NHS	Self-management of HPN versus usual care	66-year old with HPN	Lifetime; 3.5% for both	Self-management of HPN was CE ICER £1,624/QALY	Well Stroke MI Angina HF Death	CUA QALYs
Stevanovic et al, 2014[14]	CE of lowering BP in patients with HPN and low CVD risk	Dutch HIS	Anti-HPN with HCTZ versus various combinations of HCTZ/Losartan (ACEIs) or HCTZ/ARBs versus no-treatment	Various age groups: 40, 50, 60 and 65, gender and various HPN groups	10 year and lifetime; 4 % for costs and 1.5 % for health	Systolic BP reduction was found CE A 65-year old: -10 year lifetime: HCT €6,032/LYG man or €12,345/LYG woman; -Lifetime: HCT €3,076/LYG man or €3,074/LYG woman	Disease free-HPN Acute CVD (non-fatal) Stable CVD (non-fatal) Fatal CVD Non-CVD death	CEA LYG
Wu et al, 2013 [15]	CE of Amlodipine (CCB) versus ARB in the prevention of stroke and MI	Chinese Third party payer	Amlodipine (CCB) versus ARB	Average 65-year old cohort presenting HPN	5 years, 3% for both	Amlodipine was the dominant strategy	Disease free-HPN Stroke Post-stroke MI Post-MI Dead	CUA QALYs

Kourlaba et al, 2013[16]	CE of a BP lowering drug therapy in patients with mild-to-moderate HPN	Greek Third party payer	Telmisartan /HCTZ compared to Losartan/HCTZ and Valsartan/HCTZ	Average 57-year old cohort presenting HPN; analyses undertaken by gender	Lifetime, 3.5% for both	Telmisartan found to be CE Males: €3,002/QALY or €1,765/LYG Females: €10,856/QALY or €7,076/LYG	Disease free-HPN Non-fatal MI post- non-fatal MI Stroke Post-stroke death	CEA CUA LYG QALYs
Ekwunife et al, 2013 [17]	CE of drugs in the management of HPN	Nigerian third party payer perspective for costs	4 classes of antihypertensive medications: HCTZ, pranolol (Beta Blocker), lisinopril (ACE) and nifedipine (CCB)	Average 40-year olds with HPN	30 years; 3% for both	In the low CVRS ACEI had highest (15,000 $/QALY) NMB, however in the medium and high risk CVRSs, CCB had highest WTP (15,000 and 12,500 $/QALY respectively)	Non-asymptomatic (disease free) Stroke Non-fatal Stroke CHD non-fatal CHD	CUA NMB EVPI US$/ QALYs
Wisloff et al, 2012 [18]	CE of various generic anti-HPN in the prevention of CVD	Norwegian HIS	CCB compared to no-treatment in various age groups and gender	HPN patients at different age groups (40, 50, 60 and 70)	Lifetime, 4% for both	CCB / male was CE: 40: -€456,838/LYG 50: -€445,018/LYG 60: -€410,510/LYG 70: -€352,875/LYG CCB /female was CE: 40: -€621,537/LYG 50: -€630,144/LYG 60: -€588,999/LYG 70: -€465,906/LYG	Disease free-HPN Stroke Stroke-Sequelae AMI Angina HF Post-CVD Dead	CEA LYG NHB
Baker et al, 2012 [22]	CE of initiating hypertension treatment with valsartan and then switching to generic losartan in the prevention of CVD	US third party payer perspective	Two comparative analyses: 1.Continual Valsartan vs continual Losartan 2.Continual Valsartan vs Valsartan switch to generic Losartan	Moderate HPN patients – SBP 160-179- aged 18 and older	20-year time horizon and 3% discount for both	Treatment of moderate hypertension was considered CE with an ICER of $32,313/QALY or $27,268/LYG; Switching treatment resulted in an ICER of £30,170/QALY and $25,460/LYG	CVD event free with treated HPN Post-CVD with treated HPN Death	CVD event rates per arm CEA CUA LYG QALYs

Granstrom et al, 2012 [21]	Long-term CE of Candesartan versus Losartan in the primary prevention of HPN	Swedish HIS	Candesartan versus Losartan	Average 62-year old cohort presenting with HPN	Lifetime, 3% for both	Candesartan was the dominant strategy	Disease free-HPN HF Chronic IHD Post-MI PAD Post-stroke Arrhythmia Dead	CUA QALYs
Perman et al, 2011 [13]	CE multi-intervention programme versus pharmacological strategy	Argentinian Third party payer	HPN programme compared to usual care	Two target groups: 65-year old plus HPN; 65-year old, HPN and previous CVD	Lifetime, 5% for both	The HPN programme was cost-effective US$1,124/LYG	Acute myocardial event No event Death	CEA LYG
Ekman et al, 2008 [19]	CE of Irbesartan in combination with HCTZ in BP reduction	Swedish third party payer	Four strategies in male and female population: Irbesartan Placebo Losartan Valsartan	55-year old male cohort presenting with HPN	Lifetime; 3% for both	Irbesartan was CE when compared to placebo in males and females; ICERs of €3,451/QALY and €7,704/QALY respectively Losartan & Valsartan were dominated by Irbesartan in males and females	Disease free-HPN Angina MI Post-MI CHF Stroke Post-Stroke Dead	CUA QALYs
Gandjour et al, 2007 [10]	CE of a national HPN programme for patients with essential HPN and without CVD	German HIS	National programme versus no programme (for low and high risk population)	Various age groups (40-49, 50-59, and 60-69); patients with essential HPN and without CVD	Lifetime; 3% for both	National programme is CE High risk male, 40: €800/QALY, 50: €880/QALY 60: €757/QALY High risk female, 40: -€17,347/QALY 50: -€26,987/QALY 60: -€1,263/ QALY	Disease free-HPN MI Stroke Renal-failure death	CUA QALYs
Montgomery et	Effectiveness	UK Health	Anti HPN treatment	Various pop	Lifetime;	Treatment found more CE	Untreated	CUA

al, 2003 [12]	and CE of LT BP lowering	service perspective	versus non-treatment	cohorts: 30-39; 40-49; 50-59; 60-69; 70-79. Hypertensive population	6% for costs and 1.5% for effects	than non-treatment. ICER was higher for low risk women compared to low risk men	Treated T_se U_cve_ua T_cve_ua T_se_cve_ua U_cve_af T_cve_af T_se_cve_af Death	QALYs
Nordmann et al, 2003 [20]	CE of ACE as HPN first-line therapy versus conventional therapy	Canadian Third party payer	4 strategies: -Control or conventional therapy -ECG - EchoCar -ACE	40-year old male cohort presenting with HPN but without CVD	Lifetime; 5% for both	Unfavourable results of CE: ECG versus Control: US$ 0 /QALY/LYG; EchoCar vs Control: US$ 200,000/QALY/LYG ACE vs Control = US$700,000/QALY or US$525,000/LYG	Disease free-HPN (with or without LVH) CAD CVD CHF Dead	CEA CUA LYG QALYs

List of abbreviations:

ACE = Angiotensin-Converting-Enzyme Inhibitor
AMI = Acute Myocardial Infarction
ARB = Angiotensin-II-Receptor Blocker
BP = Blood pressure
CCB = Calcium-Channel Blocker
CE = cost-effectiveness or cost-effective
CEA = Cost-effectiveness analysis
CHF = Congestive Heart Failure
CUA = Cost-utility analysis
CVD = Cardiovascular disease
EchoCar = Echocardiography
EVPI = Expected Value of Perfect Information
HCTZ = Hydrochlorothiazide
HF = Heart Failure
HIS = Health Insurance System
HPN = Hypertension

IHD	= Ischaemic Heart Disease
LYG	= Life Year Gained
MI	= Myocardial infarction
NHB	= Net Health Benefit
NMB	= Net Monetary Benefit
LT	= Long term
PAD	= Peripheral Artery Disease
T_se	= Treated, side-effects (health state)
U_cve_ua	= Untreated, cardiovascular event, unaffected (health state)
T_cve_ua	= Treated, cardiovascular event, unaffected (health state)
T_se_cve_ua	= Treated, side-effects, cardiovascular event, unaffected (health state)
U_cve_af	= Untreated, cardiovascular event, affected (health state)
T_cve_af	= Treated, cardiovascular event, affected (health state)
T_se_cve_af	= Treated, side-effect, cardiovascular event, affected (health state)

Table 2 Adherence to good practice guidelines – summary results of assessment

Dimension 1 Problem concept

Information	Review question	Kaambwa et al, 2014	Stevanovic et al, 2014	Wu et al, 2013	Kourlaba et al, 2013	Ekwunife et al, 2013	Wisloff et al, 2012
Decision problem	Is there a written decision problem?	Yes	Yes	Yes	Yes	Yes	Yes
	Are the study's objective(s) consistent with the decision problem and the study's scope?	Yes	Yes	Yes	Yes	Yes	Yes
Analytical perspective	Has the perspective being stated?	Yes	Yes	Yes	Yes	Yes	Yes
Target population	Has the target population being identified?	Yes	Yes	Yes	Yes	Yes	Yes
Health outcomes	Are model's outcome(s) consistent with the perspective, scope and model's objective(s)?	Yes	Yes	Yes	Yes	Yes	Yes
	Have any adverse effect(s) be captured?	No	No	Yes	No	No	No
Interventions modelled	Are the options under evaluation clear?	Yes	Yes	Yes	Yes	Yes	Yes
	Were the inclusion/ exclusion of feasible options justified?	Yes	Yes	Yes	Yes	Yes	Yes
Time horizon	Is it sufficient to reflect all important differences between options?	Yes	Yes	No	Yes	Yes	Yes
	Have time horizon, duration of the treatment and the treatment effect(s) described and justified?	Yes	Yes	No	Yes	Yes	Yes

Continuation

Information	Review question	Baker et al, 2012 [22]	Granstrom et al, 2012	Perman G et al, 2011	Ekman et al, 2008	Gandjour et al, 2007	Montgomery et al, 2003	Nordmann A.J et al, 2003
Decision problem	Is there a written decision problem?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Are the study's objective(s) consistent with the decision problem and the study's scope?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Analytical perspective	Has the perspective being stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Target population	Has the target population being identified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health outcomes	Are model's outcome(s) consistent with the perspective, scope and model's objective(s)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Have any adverse effect(s) be captured?	No	No	No	No	No	No	No
Interventions modelled	Are the options under evaluation clear?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were the inclusion/ exclusion of feasible options justified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time horizon	Is it sufficient to reflect all important differences between options?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Have time horizon, duration of the treatment and the treatment effect(s) described and justified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dimension 2 Model concept

Information	Review question	Kaambwa et al, 2014	Stevanovic et al, 2014	Wu et al, 2013	Kourlaba et al, 2013	Ekwunife et al, 2013	Wisloff et al, 2012
Choice of model type	Was the unit of representation given?	Yes	Yes	Yes	Yes	Yes	Yes
	Does interaction(s) among individuals need to be model? If yes, was this described?	No	No	No	No	No	No
	Does the decision problem require a short time horizon?	No	No	No	No	No	No
	Is it necessary to model time in discrete cycles?	Yes	Yes	Yes	Yes	Yes	Yes
	Was a type of model discussed and chosen?	Yes	Yes	Yes	Yes	Yes	Yes
Model structure	Was the starting cohort defined by demographic and clinical characteristics affecting transition probabilities or state values?	Yes	Yes	Yes	Yes	Yes	Yes
	Were health states and transitions reflecting the biological or theoretical understanding of the disease modelled?	Yes	Yes	Yes	Yes	Yes	Yes

Continuation

Information	Review question	Baker et al, 2012 [22]	Granstrom et al, 2012	Perman et al, 2011	Ekman et al, 2008	Gandjour et al, 2007	Montgomery et al, 2003	Nordmann et al, 2003
Choice of model type	Was the unit of representation given?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Does interaction(s) among individuals need to be model? If yes, was this described?	No	No	No	No	No	No	No
	Does the decision problem require a short time horizon?	No	No	No	No	No	No	No
	Is it necessary to model time in discrete cycles?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Was a type of model discussed and chosen?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Model structure	Was the starting cohort defined by demographic and clinical characteristics affecting transition probabilities or state values?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were health states and transitions reflecting the biological or theoretical understanding of the disease modelled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dimension 3 Synthesis of evidence

Information	Review question	Kaambwa et al, 2014	Stevanovic et al, 2014	Wu et al, 2013	Kourlaba et al, 2013	Ekwunife et al, 2013	Wisloff et al, 2012
Patient heterogeneity	Was patient heterogeneity required/considered?	Yes	Yes	Yes	Yes	Yes	Yes
Data sources	Were transition probabilities and intervention effects derived from representative data sources?	Yes	Yes	Yes	Yes	Yes	Yes
	Were (all) methods and assumptions used to derive the model's inputs described?	Yes	Yes	Yes	Yes	Yes	Yes
	Were parameters derived from observational studies controlled for confounding?	NA	NA	NA	NA	NA	NA
	Was data's quality discussed?	Yes	Yes	No	Yes	Yes	Yes
	If expert opinion was used, were its methods described and justified?	No	NA	NA	NA	NA	Yes
Utilities (HSUV-weights & benefits)	Are the utilities incorporated into the model appropriate?	Yes	NA	Yes	Yes	Yes	NA
	Is the source for the utility weights referenced?	Yes	NA	Yes	Yes	Yes	NA
Half cycle correction	Was the use of a half cycle correction stated?	Yes	No	No	No	No	No
Resources including costs	Were the costs used in the model justified and its sources described?	Yes	Yes	Yes	Yes	Yes	Yes
	Were discount rates reported and justified?	Yes	Yes	Yes	Yes	Yes	Yes
Communicating results	Did the report presented results using non-technical language aided by figures or tables?	Yes	Yes	Yes	Yes	Yes	Yes
Parameter precision	Were mean value(s), distribution(s), source(s) of data and rationale for the supporting evidence described?	Yes	Yes	No	Yes	Yes	Yes

Continuation

Information	Review question	Baker et al, 2012 [22]	Granstrom et al, 2012	Perman et al, 2011	Ekman et al, 2008	Gandjour et al, 2007	Montgomery et al, 2003	Nordmann et al, 2003
Patient heterogeneity	Was patient heterogeneity required/considered?	No	Yes	No	Yes	Yes	Yes	Yes
Data sources	Were transition probabilities and intervention effects derived from representative data sources?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were (all) methods and assumptions used to derive the model's inputs described?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were parameters derived from observational studies controlled for confounding?	NA	Yes	NA	NA	NA	Yes	NA
	Was data's quality discussed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	If expert opinion was used, were its methods described and justified?	NA	NA	NA	NA	No	NA	Yes
Utilities (HSUV-weights & benefits)	Are the utilities incorporated into the model appropriate?	Yes	Yes	NA	Yes	Yes	Yes	Yes
	Is the source for the utility weights referenced?	Yes	Yes	NA	Yes	Yes	Yes	Yes
Half cycle correction	Was the use of a half cycle correction stated?	No	Yes	No	No	Yes	No	Yes
Resources including costs	Were the costs used in the model justified and its sources described?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were discount rates reported and justified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Communicating results	Did the report presented results using non-technical language aided by figures or tables?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parameter precision	Were mean value(s), distribution(s), source(s) of data and rationale for the supporting evidence described?	Yes	Yes	Yes	Yes	Yes	No	No

Dimension 4 Analysis of uncertainty

Information	Review question	Kaambwa et al, 2014	Stevanovic et al, 2014	Wu et al, 2013	Kourlaba et al, 2013	Ekwunife et al, 2013	Wisloff et al, 2012
Analysis of uncertainty	Was analysis of uncertainty pertaining to the decision problem included and reported?	Yes	Yes	Yes	Yes	Yes	Yes
Parameter estimation & uncertainty	Were one-way or two-way DSA sensitivity analysis performed?	No	Yes	Yes	No	No	No
	Was a probabilistic sensitivity analysis (PSA) included?	Yes	Yes	No	Yes	Yes	Yes
Multivariate estimation and correlation	Was correlation among parameters considered?	NA	NA	NA	NA	NA	NA
Structural uncertainty	Were there any discussion /evidence of uncertainty in structural assumptions?	Yes	Yes	No	No	No	Yes
Other reporting of uncertainty analyses	Was EVPI measured/ discussed?	No	No	No	No	Yes	No
	If model calibration was used to estimate parameters, was uncertainty tested?	NA	NA	NA	NA	NA	NA

Continuation

Information	Review question	Baker et al, 2012 [22]	Granstrom et al, 2012	Perman et al, 2011	Ekman et al, 2008	Gandjour et al, 2007	Montgomery et al, 2003	Nordmann et al, 2003
Analysis of uncertainty	Was analysis of uncertainty pertaining to the decision problem included and reported?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parameter estimation & uncertainty	Were one-way or two-way DSA sensitivity analysis performed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Was a probabilistic sensitivity analysis (PSA) included?	No	Yes	Yes	No	Yes	No	No
Multivariate estimation and correlation	Was correlation among parameters considered?	NA	NA	NA	NA	NA	NA	NA
Structural uncertainty	Were there any discussion /evidence of uncertainty in structural assumptions?	No	No	Yes	Yes	Yes	No	Yes
Other reporting of uncertainty analyses	Was EVPI measured/ discussed?	No	No	No	No	No	No	No
	If model calibration was used to estimate parameters, was uncertainty tested?	NA	NA	NA	NA	NA	NA	NA

Dimension 5 Model transparency and validation

Information	Review question	Kaambwa et al, 2014	Stevanovic et al, 2014	Wu et al, 2013	Kourlaba et al, 2013	Ekwunife et al, 2013	Wisloff et al, 2012
Transparency	Were the purpose, type and graphical description of the model provided?	Yes	Yes	Yes	Yes	Yes	Yes
	Were the source(s) of funding and their role identified?	Yes	Yes	Yes	Yes	Yes	Yes
	Were data sources identified/ described?	Yes	Yes	Yes	Yes	Yes	Yes
	Were methods customised to specific application(s) and settings?	Yes	Yes	Yes	Yes	Yes	Yes
	Were the effects of uncertainty measured?	Yes	Yes	Yes	Yes	Yes	Yes
	Were limitations acknowledged/ discussed?	Yes	Yes	Yes	Yes	Yes	Yes
	Was any reference to the availability of model's documentation at request or terms and conditions to access it?	No	No	No	No	No	No
Validation	Was there any evidence of model's face validity?	Yes	Yes	Yes	Yes	Yes	Yes
	Was internal validity (verification or technical validity) assessed?	Yes	Yes	Yes	Yes	Yes	Yes
	Was cross-validation (external consistency) assessed?	Yes	Yes	No	Yes	Yes	No
	Was external validity assessed	No	Yes	No	No	No	No
	Was the model's predictive validity assessed?	NA	NA	NA	NA	NA	NA

Continuation

Information	Review question	Baker et al, 2012 [22]	Granstrom et al, 2012	Perman et al, 2011	Ekman et al, 2008	Gandjour et al, 2007	Montgomery et al, 2003	Nordmann et al, 2003
Transparency	Were the purpose, type and graphical description of the model provided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were the source(s) of funding and their role identified?	Yes	Yes	No	Yes	No	Yes	Yes
	Were data sources identified/ described?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were methods customised to specific application(s) and settings?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were the effects of uncertainty measured?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were limitations acknowledged/ discussed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Was any reference to the availability of model's documentation at request or terms and conditions to access it?	No	No	No	No	No	No	No
Validation	Was there any evidence of model's face validity?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Was internal validity (verification or technical validity) assessed?	Yes	Yes	Yes	Yes	Yes	Yes	No
	Was cross-validation (external consistency) assessed?	No	No	Yes	Yes	Yes	Yes	Yes
	Was external validity assessed	No	No	No	No	No	No	No
	Was the model's predictive validity assessed?	NA	NA	NA	NA	NA	NA	NA

SUPPLEMENTAL MATERIAL

Appendix 1

Search strategy: Cochrane databases (searched 20 March 2015 14:00:51,676 for the period 2000 to 2015)

NHS EED (economic evaluations)

ID	Searches - CRD (NHS-EED)
#1	MeSH blood pressure EXPLODE PERMUTE
#2	MeSH hypertension EXPLODE PERMUTE
#3	cost utility analys*
#4	mathematical model
#5	decision analys*
#6	Markov chain* or Markov process* or decision tree
#7	Economics
#8	cost effective* or cost effective* analys*
#9	#3 OR #4 OR #5 OR #6 OR #7 OR #8
#10	#1 OR #2
#11	#9 AND #10
#12	MeSH primary prevention EXPLODE PERMUTE
#13	#11 AND #12

NHS EED National Health Service Economic Evaluation Database, */$* wildcard characters

Appendix 2

EMBASE and MEDLINE databases (searched 20 March 2015 16:59) via OVID MEDLINE(R)

ID	Searches (via OVID)
#1	(lowering blood pressure or lowering-blood-pressure or blood pressure lowering).mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, an, ui]
#2	(hypertensi$ or antihypertensi$ or anti-hypertensi$).mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, an, ui]
#3	1 OR 2
#4	(cost effective$ OR cost-effective$ OR mathematical model OR decision-analys$s OR decision analys$s OR Markov OR decision tree OR economic evaluation OR cost utility).mp. [mp=ti, ab, sh, hw, tn, ot, dm, mf, dv, kw, nm, kf, px, rx, an, ui]
#5	3 AND 4
#6	limit 5 to English language
#7	limit 6 to yr="2000 -Current"
#8	limit 7 to humans
#9	Exclude conference abstracts, methodological papers, commentaries, editorials, notes
#10	remove duplicates from 9

Appendix 3 Framework to assess adherence to good practice guidelines in Decision-Analytic Modelling (DAM)

Source: Peñaloza et al. *A Systematic Review of Research Guidelines in Decision-Analytic Modeling*. Value in Health 18 (2015), Table 5, p. 524-527.

DIMENSION 1: PROBLEM CONCEPT			
Components of good practice	Questions for review	Yes, No, or NA	Attributes
Decision problem	Is there a written statement of the decision problem and scope of the study?		A clear statement of the decision problem and scope would determine the interventions and health outcomes to be measured
	Are the objective(s) of the study and model structure consistent with the stated decision problem and scope?		They are expected to be consistent
Analytical perspective	Has the perspective of the model been stated?		Most common perspectives are: patient, health system (insurer) and society
Target population	Has the target population been identified?		Target population should be defined in terms of features relevant to the decision (geography, patient characteristics, including comorbid conditions, disease prevalence and stage)
Health outcomes	Are the outcomes of the model stated and consistent with the perspective, scope and overall objective(s) of the model?		Health outcomes may be events, cases of disease, deaths, life-years gained, quality-adjusted life-years, disability-adjusted life-years or other measures important to stakeholders and should be directly relevant to the question being asked
	Has any adverse effect of the intervention(s) been captured?		Interventions may cause negative health consequences that need to be modelled and discussed as part of the study's results. The impact of assumptions regarding adverse effects of interventions should be assessed as part of the structural uncertainty analysis
Comparators	Is there a clear definition of the alternative interventions under evaluation?		Usually the choice of comparators is governed by the scope of the model. Impact of assumptions adopted when deciding upon comparators should be assessed as part of the structural uncertainty analysis
	Is there a discussion around feasible options or justification for the exclusion of feasible options?		The choice of comparators affects results and should be determined by the decision problem, not by data availability. All feasible and practical strategies as determined by the scope of the model should be considered. Constraining the range of strategies should be justified
Time horizon	Is the time horizon of the model justified and sufficient to reflect all important differences between options?		Time horizon of the model should be long enough to capture relevant differences in outcomes across strategies (lifetime). Time horizon is dictated by the problem scope

Note: NA= Not Apply

DIMENSION 2: MODEL CONCEPT			
Components of good practice	Questions for review	Yes, No, or NA	Attributes
Choice of model type	Has the unit of representation been given?		Usually stated in terms of groups or individuals. If groups are being modelled most frequently decision trees, Markov processes or infectious disease models are the correct choice; if individuals are being modelled then the choice is between DES, dynamic transmission models or agent-based models
	Is there a need to model the interaction between individuals in this model? Has this been discussed?		If interactions between individuals is required (when the disease or treatment includes interactions between individuals) then DES, dynamic-transmission, or agent-based models may be the correct choice
	Does the decision problem require a short time horizon?		For simple models or problems (short time horizon, few outcomes) a decision tree may be appropriate; time horizon should be large enough to capture all health effects and costs directed related to the decision problem
	Is it necessary to model time in discrete cycles?		Continuously for Individual STM or in discrete cycles for Markov STM; if the assumption that transition probabilities do not depend on history is not required, then individual state-transition models are an alternative; If disease or treatment process need to be represented as health states, state transition models are appropriate (Markov type)
	Is there a need to model competition for resources or the development of waiting lists or queues?		If the problem requires the ability of a model to incorporate interactions between individuals and other model parts for example to answer questions on resource allocation i.e., organ allocation for transplantation, distribution of antiretroviral medications in resource-poor environments, then a DES may be appropriate
	Has a type of model been chosen and discussed?		It is expected that studies report on the reasons for choosing a type of model
Model structure	Has the starting cohort been defined by demographic and clinical characteristics affecting the transition probabilities or state values?		If results may vary by subgroups (age, sex, risk factors) is advisable to report results for different cohorts
	Has health states and transitions reflecting the biological/theoretical understanding of the disease or condition been modelled?		States should adequately capture the type of intervention (prevention, screening, diagnostics, and treatment) as well as the intervention's benefits and harms. States need to be homogeneous with respect to both observed and unobserved characteristics that affect transition probabilities

Note: NA= Not Apply

DIMENSION 3: SYNTHESIS OF EVIDENCE			
Components of good practice	Questions for review	Yes, No, or NA	Attributes
Data sources	Has transition probabilities and intervention effects been derived from representative data sources for the decision problem?		Most common sources of data include population-based epidemiological studies, control arms of trials or literature
	Has (all) methods and assumptions used to derive transition probabilities and intervention effects been described/justified?		Attention should be given to the use of transition probabilities and rates; conversion of transition probabilities from one time unit to another should be done through rates and never presented as percentages
	Has parameters relating to the effectiveness of interventions derived from observational studies been controlled for confounding?		If results of meta-analyses were used as data sources then consider how potential confounders are addressed; consider the likelihood of increased heterogeneity resulting from residual confounding and from other biases across studies. Efficacy derived from RCT may have to be adjusted for compliance to reflect real-world effectiveness. Effectiveness derived from observational studies must be adjusted for confounding (e.g., using multivariate regression techniques or propensity scoring). Adjustment for time-varying confounding (confounders that simultaneously act as intermediate steps in the pathway between intervention and outcome) require special methods such as marginal structural analysis or g-estimation. When results from observational studies are used in the model, causal graphs can be used to explicitly state causal assumptions
	Has the quality of the data been assessed appropriately?		Sources of data and data limitations are expected to be discussed
	Has expert opinion been used, are the methods described and justified?		An expectation that strengths and limitations of assumptions adopted should be included
Utilities	Are the utilities incorporated into the model appropriate?		methods used to obtain utility weights and methodology used to transform health estate estimates into quality of life scores
	Is the source for the utility weights referenced?		Sources of data and data limitations are expected to be discussed
Cycle length and half cycle correction	Has the choice of cycle length been justified?		It should be based on the clinical problem and remaining life expectancy
	Has the use of a half cycle correction been stated?		Any assumption adopted is expected to be disclosed
Resources/ costs	Are the costs incorporated into the model justified and sources described?		Sources of data and data limitations are expected to be discussed
	Has discount rates been reported and justified given the target decision-maker?		
Patient heterogeneity	Has patient heterogeneity been considered?		For example, in a cohort model states need to be homogeneous to observed or unobserved characteristics affecting transition probabilities to observed or unobserved characteristics affecting transition probabilities
Parameter precision	Has mean values and distributions around the mean and the source and rationale for the supporting evidence been clearly described for ***each parameter*** included in the model?		Sources of data and data limitations are expected to be discussed

DIMENSION 4: ANALYSIS OF MODEL UNCERTAINTY			
Components of good practice	Questions for review	Yes, No, or NA	Attributes
Uncertainty	Has analyses of uncertainty pertaining to the decision problem been included and reported? If not, has the reasons been explained for its omission?		Analysis of uncertainty is expected to be include as part of the DAM
Parameter estimation & uncertainty	Has one-way DSA or two-way sensitivity analysis been performed?		Tornado diagrams, threshold plots or simple statements of threshold parameter values, are all appropriate. Uncertainty of parameters may be represented by several discrete values, instead of a continuous range, called 'scenario analyses'. It is a good practice to include the specification of parameter's point estimate and a 95% CI range.
	Has a Probabilistic Sensitivity Analysis (PSA) been included?		The specific distribution (e.g. Beta, normal, lognormal) as well as its parameters should be disclosed. When PSA is performed without an accompanying EVPI, options for presenting results include CEAC and distributions of net monetary benefit or net health benefit. When more than two comparators are involved, curves for each comparator should be plotted on the same graph.
	Has correlation among parameters been assessed?		Lack of evidence on correlation among parameters should not lead to an assumption of independence among parameters
	If model calibration was used to derive parameters, has the uncertainty around calibrated values been tested using DSA or PSA?		Calibration is commonly used to estimate parameters or adjust estimated values such as overall and disease specific mortality and event incidence rates
Structural uncertainty	Has a discussion about the inclusion/exclusion of assumptions affecting the structure of the model been included? (refers to potentially relevant comparators, health states and recurrent events or any other assumption affecting the structure of the model)		For example: i) health states and the strategies adopted following the recurrence of events; ii) length of treatment effects; iii) types of adverse effects included; iv) duration of treatment effects; v) time dependency of probabilities (in a time dependent utility, the cost of delaying treatment as a function of the time a patient has remained in an untreated acute pathological state); vi) prognostic implications of surrogate end points; vii) clinical events; viii) comparators. Although these structural assumptions are not typically quantified, it is uncertain whether they express reality accurately and for that reason they should be assessed as part of structural uncertainty analysis
Other reporting of uncertainty analyses	Has the EVPI being measured /discussed?		If the purpose of a PSA is to guide decisions about acquisition of information to reduce uncertainty in the results, EVPI should be presented in terms of expected value of information. EVPI is commonly reported in monetary terms using net monetary benefit or net health benefits; EVPI should be reported for specified ICER thresholds

Note: NA= Not Apply

DIMENSION 5: MODEL TRANSPARENCY AND VALIDATION			
Components of good	Questions for review	Yes,	Attributes

practice		No, or NA	
Transparency	Has a graphical description of the model been provided?		
	Has all sources of funding and their role been identified?		
	Has all methods used been customised to specific application(s) and settings?		
	Has the report used nontechnical language and clear figures and tables to enhance the understanding of the model?		
	Has limitations and strengths been acknowledged/discussed?		
	Is there any reference as to whether technical documentation would be made available at request?		
Validation	Is there any evidence of model's face validity?		Can occur in several ways: the group that develop the model can appeal to members of the modelling group, people in the same organisation who did not build the model, or external consultants. Any reader can perform his/her own evaluation. Peer review (previous to publication)
	Has internal validity been assessed?		Verification or technical validity; models should be subject to rigorous verification and the methods used should be described and results made available on request
	Has cross-validation been assessed?		or external consistency (involves examining different models that address the same problem and comparing their results) its meaningfulness depends on the degree to which methods and data are independent. Modellers should search for modelling analyses of the same or similar problems and discuss insights gained from similarities and differences in results
	Has external validity been assessed?		This compares the model's results with actual event data; a formal process needs to be developed including identifying suitable sources of data; results of external validation should be made available
	Has the model's predictive validity been assessed?		If feasible given the decision problem and future's sources availability

Note: NA= Not Apply

References

1. WHO. *GLOBAL STATUS REPORT on noncommunicable diseases 2014*. Geneva2014.
2. NICE. Lipid modification: cardiovascular risk assessment and the primary and secondary prevention of cardiovascular disease. NICE Guidelines CG67. London: National Institute for Health and Clinical Excellence, 2008.
3. NICE. *Hypertension: clinical management of primary hypertension in adults. NICE clinical guideline 127*. London: National Institute for Health and Clinical Excellence, 2011.
4. Gray AM. Cost-effectiveness analyses alongside randomised clinical trials. *Clin Trials*. 2006; 3: 538-42.
5. Ramsey S, Willke R, Briggs A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health*. 2005; 8: 521-33.
6. Peñaloza Ramos MC, Barton P, Jowett S and Sutton AJ. A Systematic Review of Research Guidelines in Decision-Analytic Modeling. *Value in Health*. 2015; 18: 512-29.
7. Moher D, Liberati A, Tetzlaff J and Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA StatementThe PRISMA Statement. *Annals of Internal Medicine*. 2009; 151: 264-9.
8. Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision analytic modeling in health-care evaluation: Report of the ISPOR task force on good research practices - Modeling studies. *Value in Health*. 2003; 6: January/February.
9. Philips Z, Ginnelly L, Sculpher M, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health technology assessment (Winchester, England)*. 2004; 8: Sep.
10. Gandjour A and Stock S. A national hypertension treatment program in Germany and its estimated impact on costs, life expectancy, and cost-effectiveness. *Health Policy*. 2007; 83: 257-67.
11. Kaambwa B, Bryan S, Jowett S, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): A cost-effectiveness analysis. *European Journal of Preventive Cardiology*. 2014; 21: 1517-30.
12. Montgomery AA, Fahey T, Ben-Shlomo Y and Harding J. The influence of absolute cardiovascular risk, patient utilities, and costs on the decision to treat hypertension: A Markov decision analysis. *Journal of Hypertension*. 2003; 21: 1753-9.
13. Perman G, Rossi E, Waisman GD, et al. Cost-effectiveness of a hypertension management programme in an elderly population: A Markov model. *Cost Effectiveness and Resource Allocation*. 2011; 9.
14. Stevanovic J, O'Prinsen AC, Verheggen BG, Schuiling-Veninga N, Postma MJ and Pechlivanoglou P. Economic evaluation of primary prevention of cardiovascular diseases in mild hypertension: A scenario analysis for the Netherlands. *Clinical Therapeutics*. 2014; 36: 368-84.e5.
15. Wu Y, Zhou Q, Xuan J, et al. A cost-effectiveness analysis between amlodipine and angiotensin ii receptor blockers in stroke and myocardial infarction prevention among hypertension patients in China. *Value in Health Regional Issues*. 2013; 2: 75-80.
16. Kourlaba G, Fragoulakis V, Theodoratou D and Maniadakis N. Economic evaluation of telmisartan, valsartan and losartan in combination with hydrochlorothiazide for treatment of mild-to-moderate hypertension in Greece: A cost-utility analysis. *Journal of Pharmaceutical Health Services Research*. 2013; 4: 81-8.
17. Ekwunife OI, Okafor CE, Ezenduka CC and Udeogaranya PO. Cost-utility analysis of antihypertensive medications in Nigeria: A decision analysis. *Cost Effectiveness and Resource Allocation*. 2013; 11.
18. Wisloff T, Selmer RM, Halvorsen S, Fretheim A, Norheim OF and Kristiansen IS. Choice of generic antihypertensive drugs for the primary prevention of cardiovascular disease - A cost-effectiveness analysis. *BMC Cardiovascular Disorders*. 2012; 12.

19. Ekman M, Bienfait-Beuzon C and Jackson J. Cost-effectiveness of irbesartan/hydrochlorothiazide in patients with hypertension: An economic evaluation for Sweden. *Journal of Human Hypertension*. 2008; 22: 845-55.
20. Nordmann AJ, Krahn M, Logan AG, Naglie G and Detsky AS. The cost effectiveness of ACE inhibitors as first-line antihypertensive therapy. *PharmacoEconomics*. 2003; 21: 573-85.
21. Granstrom O, Levin L and Henriksson M. Cost-effectiveness of candesartan versus losartan in the primary preventive treatment of hypertension. *ClinicoEconomics and Outcomes Research*. 2012; 4: 313-22.
22. Baker TM, Goh J, Johnston A, Falvey H, Brede Y and Brown RE. Cost-effectiveness analysis of valsartan versus losartan and the effect of switching. *Journal of medical economics*. 2012; 15: 253-60.
23. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P and Krahn M. Conceptualizing a Model: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. *Value in Health*. 2012; 15: 804-11.
24. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM and Wong JB. Model Transparency and Validation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. *Value in Health*. 2012; 15: 843-50.
25. Jackson CH, Bojke L, Thompson SG, Claxton K and Sharples LD. A framework for addressing structural uncertainty in decision models. *Medical Decision Making*. 2011; 31: 662-74.
26. Squires H, Chilcott J, Akehurst R, Burr J and Kelly MP. A systematic literature review of the key challenges for developing the structure of public health economic models. *International journal of public health*. 2016; 61: 289-98.
27. Briggs AH. Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics*. 2000; 17: 479-500.
28. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ and Paltiel AD. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value in Health*. 2012; 15: 835-42.
29. Caro JJ and Moller J. Decision-analytic models: current methodological challenges. *Pharmacoeconomics*. 2014; 32: 943-50.
30. Squires H, Chilcott J, Akehurst R, Burr J and Kelly MP. A Framework for Developing the Structure of Public Health Economic Models. *Value in Health*.